Mathematical Modeling and Analysis

Los Alamos NATIONAL LABORATORY EST.1943

Introduction

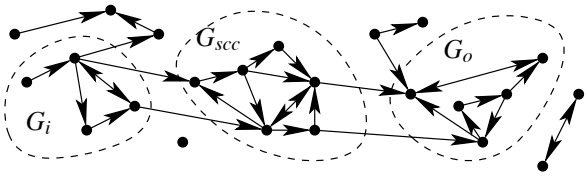
Infectious diseases spread along human, animal, plant, or computer networks. Understanding properties of these networks gives insight into analyzing how the disease spreads, and so recent work has investigated the impact of network properties on epidemic spread. [1, 3, 4]. Most studies assume an average *transmissibility* (the probability that an infection of node i will result in infection of the neighboring node j), but it is well-known that the transmissibility is heterogeneously distributed. In this study we investigate the impact of this heterogeneity and derive rigorous bounds on the size and probability of epidemics for given average transmissibilities. The bounds we find give insight into optimal strategies to prevent or reduce the impact of epidemics.

We study epidemics spreading on large random networks, with fixed degree distribution. The nodes are divided into three classes:

- *Susceptible*: Nodes which may become infected if a neighbor is infected.
- *Infected*: Nodes which are infected and may infect susceptible neighbors.
- *Recovered*: Nodes which have been infected but are no longer. These nodes may not infect or be infected.

The infectiousness of an individual can depend on a number of properties: for example, the levels of virus shedding or whether an employer allows sick days. Similarly the susceptibility of an individual may depend on vaccination or exposure history as well as personal protective measures. The transmissibility T_{uv} from node u to v is given by $T(\mathcal{I}_u, \mathcal{S}_v)$ where $\mathcal{I}_u$ and $\mathcal{S}_v$ represent all factors affecting the infectiousness of u and susceptibility of v respectively.

A standard approach to epidemic modeling is to take a single infected individual u (the *index case*) and consider its neighbors. Each neighbor v is infected with probability T_{uv}. The index case then recovers. We then consider the newly infected nodes and their susceptible neighbors, repeating until no infected nodes remain.



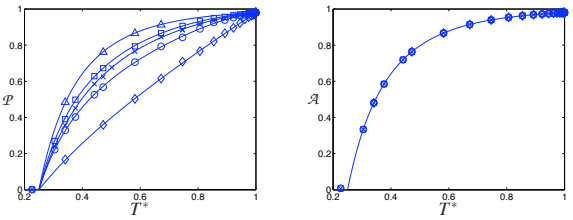
Schematic representation of G_i, G_{scc}, and G_o. All nodes in G_{scc} can reach any other node in G_{scc}.

An equivalent approach is more computationally intense, but provides a useful theoretical framework. We consider each node u separately and determine *a priori* whether u would infect its neighbor v if u becomes infected while v is susceptible. If so, we place a directed edge from u to v. The edges of the original network are either lost or replaced with a directed edge, which may be bidirectional. This is represented in figure 1. The index case is then chosen, and all nodes in its out-component are infected. The size of the outbreak is equal to the out-component size. If the average transmissibility is large enough, then some of the nodes have an out-component which is limited in size only by the size of the network, that is, G_i (the giant in-component), G_{scc} (the giant strongly-connected component) and G_o (the giant out component) exist as shown in figure 1. We define these large outbreaks to be *epidemics*. This occurs whenever the initial infection is in either G_i or G_{scc}. The size of the epidemic is equal to that of G_{scc} and G_o combined. Thus the probability $\mathcal{P}$ of an epidemic is given by the fraction of the nodes in $G_i \cup G_{scc}$, while the attack rate $\mathcal{A}$ (the fraction infected) is given at leading order by the fraction of nodes in $G_{scc} \cup G_o$. We note that a symmetry exists interchanging these by interchanging the direction of the arrows, so any method that calculates the probability of an epidemic may be modified to calculate the size.

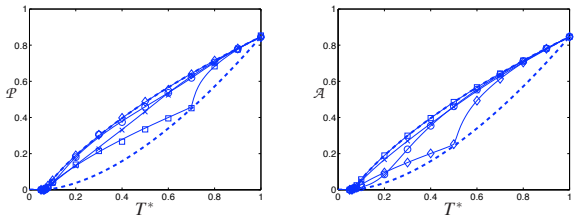
Calculating Epidemic Probability

The random networks we study have few short loops in the limit of large network size, and so early in the outbreak we may assume the outbreak spreads on an infinite tree. We use a generating

Predicting the size and probability of epidemics in a population with heterogeneous infectiousness and susceptibility



Comparison of theory (lines) with simulation (symbols) for an Erdős–Rényi network. For the different distributions of infectivity (with susceptibility constant), $\mathcal{P}$ changes, but $\mathcal{A}$ does not. We use constant recovery time $\tau = 5$ ($\triangle$), $\tau = 0$ or ∞ ($\Diamond$), $\tau = 2$ or 8 ($\square$), $\tau = 1$ or 10 ($\circ$), and finally a constant recovery rate ($\times$).



Comparison of theory (curves) with simulation (symbols) for $T_{uv} = 1 - \exp(-\alpha I_u S_v)$ in a scale-free network with a cutoff at high degree. The theoretical bounds are in dashed bold. The distributions are $\Diamond$: $P(I) = \delta(I-1)$, $P(S) = 0.5\delta(S - 0.001) + 0.5\delta(S-1)$; $\times$: $P(I) = 0.5\delta(I-0.1) + 0.5\delta(I-1)$, $P(S) = 0.2\delta(S-0.1) + 0.8\delta(S-1)$; $\circ$: $P(I) = 0.5\delta(I-0.1) + 0.5\delta(I-1)$, $P(S) = 0.8\delta(S-0.011) + 0.2\delta(S-1)$; $\square$: $P(I) = 0.3\delta(I-0.001) + 0.7\delta(I-1)$, $P(S) = \delta(S-1)$.

function [5] approach where a probability distribution of non-negative integers k may be encoded as a function by $f(x) = \sum_{k=0}^{\infty} p_k x^k$. We will calculate a generating function $f(x,g)$ for the number of infected individuals in generation g. The probability that the epidemic dies out by generation g is given by $f(0,g)$. The probability that the epidemic dies in the limit of infinite system size is given by $\lim_{g\to\infty} f(0,g)$.

The details of the calculation are slightly technical, but can be found in [2]. We find rigorous upper and lower bounds on the probability of an epidemic in terms of the average transmissibility. For a fixed average transmissibility T_0, the epidemic is most likely if all nodes have probability T_0 to infect a random neighbor. In contrast, the epidemic is least likely if a fraction T_0 of the nodes will infect all their neighbors while the remaining nodes infect none. Equivalent statements hold for susceptibility and epidemic size. Our results hold for Erdős–Rényi and scale-free networks as shown in the figures.

Discussion

The bounds we have derived show that if we maximize the variance in infectiousness for a given average transmissibility, we minimize the probability of an epidemic. Similarly if we maximize the variance in susceptibility, we minimize the size of an epidemic. Consider two strategies that have the same average effect, but one has heterogeneous impact on infectiousness (*e.g.*, incomplete contact tracing) while the other has heterogeneous impact on susceptibility (*e.g.*, incomplete vaccination). Which is optimal will depend on whether an outbreak is established or not. If it is not established and the strategy is intended to prevent an epidemic, then contact-tracing will outperform vaccination. In contrast, if it is established and we instead hope to reduce the size, vaccination will outperform contact tracing.

References

[1] Lauren Ancel Meyers, Mark Newman, and B. Pourbohloul. Predicting epidemics on directed contact networks. *Journal of Theoretical Biology*, 240(3):400–418, June 2006.

[2] Joel C. Miller. Predicting the size and probability of epidemics in a population with heterogeneous infectiousness and susceptibility. *ArXiv/q-bio/0702007*, 2007.

[3] Mark E. J. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66, 2002.

[4] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, Apr 2001.

[5] Herbert S. Wilf. *generatingfunctionology*. A K Peters, Ltd, 3 edition, 2005.

http://math.lanl.gov/, T-7, MS B284, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545